Identification of Cassava MicroRNAs under Abiotic Stress

Carolina Ballén-Taborda[1], Germán Plata[2], Sarah Ayling[3], Fausto Rodríguez-Zapata, Luis Augusto Becerra Lopez-Lavalle[1], Jorge Duitama[1] & Joe Tohme[1]

[1]Agrobiodiversity and Biotechnology Project. International Center for Tropical Agriculture (CIAT). A.A. 6713, Cali, 76001000, Colombia.
[2]Center for Computational Biology and Bioinformatics, Columbia University, 1130 St. Nicholas Ave. New York, NY 10032, USA.
[3]The Genome Analysis Centre. Norwich Research Park. Norwich. NR4 7UH. UK.

✉ c.ballen@cgiar.org

Introduction

MicroRNAs (miRNAs) are small non-coding endogenous RNA molecules that regulate gene expression in plants and animals. MiRNAs play a key role in multiple biological processes including stem cell differentiation, organ development, signaling, and response to biotic and abiotic stresses. Cassava (*Manihot esculenta* Crantz), one of the most important crops in tropical regions of the world, is tolerant to drought and other adverse conditions, making it a desirable model for understanding post-transcriptional control in plants in the light of climate change. Most miRNAs discovered in cassava were predicted using bioinformatics alone or through sequencing of plants challenged by biotic stress. Here, we use deep-sequencing and a combination of bioinformatics methods to identify potential cassava miRNAs expressed in different tissues subject to heat and drought conditions. We predict 881 cassava miRNAs and 1136 possible gene targets. To validate our approach, we verified the condition-specific expression of 5 cassava small RNAs using real time PCR. We also found a significantly lower expression of the predicted target genes under drought stress compared to other cassava genes. Finally, gene ontology enrichment analysis allowed us to identify several interesting miRNAs that may play a role in stress-induced post-transcriptional regulation in cassava and other plants.

Results and Discussion

Deep sequencing of cassava miRNAs:
We obtained over 14 million Illumina reads with a expected read length distribution for miRNA sequencing (Figure 2). 60 conserved sequences obtained at the end of the analysis pipeline were joined in 26 families (Figure 3) based on the names already established for corresponding miRNAs in other species.
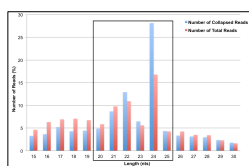

Figure 2: Frequency distribution of cassava small RNAs between 15 to 30 nt. Reads with 20 to 25 nts (black box) in length were selected for further analysis
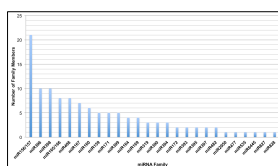
Figure 3. Distribution of 60 conserved cassava miRNAs grouped in 26 different families. Names correspond to homologous small RNAs in other plant species.

Target gene identification and functional analyses:
Targets of conserved miRNAs showed a significant enrichment of transcription factors (Figure 4A)

- MiR156 targeted genes in the squamosa promoter-binding family
- MiR159 targeted a MYB-like regulatory protein
- MiR160 which was previously associated to drought stress [2] and which we found to target several auxin response factors (AFR10, AFR16, AFR17). AFR10 and AFR16 were shown to increase their expression in the roots of drought treated *Sorghum bicolor*, suggesting a common mechanism in cassava given our observation of lower miR160 expression under those conditions.

Targets of non-conserved miRNAs were mainly enriched in protein modification enzymes involving multiple kinases and phosphatases (Figure 4B).

- Protein phosphatase homologous to an abscisic acid (ABA) induced gene in *A. thaliana* (HAI3). ABA plays important roles in abiotic stress signaling [3].
- Gibberellin oxidase which is differentially regulated under stress conditions in *Zoysia japonica* [4].
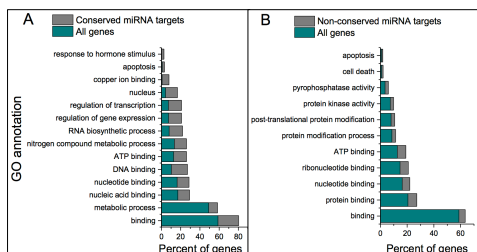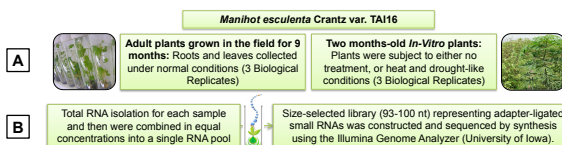

Figure 4. GO enrichment analysis of predicted miRNA targets. (A) Significantly overrepresented GO terms for conserved miRNAs identified in this study. (B) Significantly overrepresented GO terms for possible cassava-specific miRNAs. Cyan and gray bars indicate the fraction of miRNA targets and cassava genes annotated with a corresponding GO term, respectively. See supplementary tables 5 and 6 for the full list of significant terms.

Analysis Pipeline
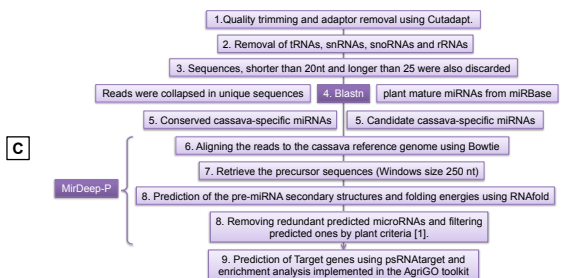
Deep sequencing of smallRNAs

A — *Manihot esculenta* Crantz var. TAI16

Adult plants grown in the field for 9 months: Roots and leaves collected under normal conditions (3 Biological Replicates)

Two months-old *In-Vitro* plants: Plants were subject to either no treatment, or heat and drought-like conditions (3 Biological Replicates)

B — Total RNA isolation for each sample and then were combined in equal concentrations into a single RNA pool → Size-selected library (93-100 nt) representing adapter-ligated small RNAs was constructed and sequenced by synthesis using the Illumina Genome Analyzer (University of Iowa).

Processing and Analysis of the cassava smallRNA sequences

C

1. Quality trimming and adaptor removal using Cutadapt.
2. Removal of tRNAs, snRNAs, snoRNAs and rRNAs
3. Sequences, shorter than 20nt and longer than 25 were also discarded

Reads were collapsed in unique sequences | 4. Blastn plant mature miRNAs from miRBase

5. Conserved cassava-specific miRNAs | 5. Candidate cassava-specific miRNAs

MirDeep-P:
6. Aligning the reads to the cassava reference genome using Bowtie
7. Retrieve the precursor sequences (Windows size 250 nt)
8. Prediction of the pre-miRNA secondary structures and folding energies using RNAfold
8. Removing redundant predicted microRNAs and filtering predicted ones by plant criteria [1].
9. Prediction of Target genes using psRNAtarget and enrichment analysis implemented in the AgriGO toolkit

Validation by qRT-PCR of cassava miRNAs related to stress response

D — 5 conserved candidate miRNAs previously associated in stress response were select for the real time PCR (Using TaqMan MicroRNA Assays. ath-miR156a, ath-miR159a, ath-miR160a, ath-miR397a and ath-miR408)
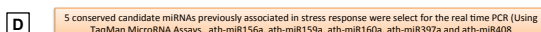
Figure 1: A. Plant materials and treatments. **B.** Total RNA Isolation and Sequencing, **C.** Workflow of the Bioinformatics analysis and **D.** Validation off conserved microRNAs by real time PCR (qRT-PCR)

Table 1: Summary of the results of the identification of cassava microRNA and their target genes.

Procedure	Number of sequences	
Total od reads	14,565,645	
1. Adaptor and quality trimming	9,570,232 Total Reads	
2. Removal of rRNA, tRNA, snRNA, snoRNAs	598,120 Collapsed Reads	
	2,920,905 Total Reads	
3. 20-25nt filtering	391,453 Collapsed Reads	
	1,553,668 Total Reads	
4. Blast against plant mature miRNAs from miRBase	Conserved	Non-Conserved
5. Total collapsed Reads	981	390,472
6. Aligned to the genome	146 (14.88%)	151,390 (38.77%)
7. Number of precursors	338	441,567
8. Validated By MirDeep-P / clusters	114 / 64	12247 / 8325
8. Precursors Validated by Meyers 2008 Filter / Unique Reads	106 / 60	1103 / 821
9. Target Genes	134	1002

Validation of cassava miRNAs related to stress response using qRT-PCR

To validate our bioinformatics analysis, we selected 5 conserved candidate miRNAs with *A. thaliana* homologs previously associated in stress response [5] (ath-miR156a, ath-miR159a, ath-miR160a, ath-miR397a, and ath-miR408). We identified precursors for each miRNA in the cassava reference genome (Figure 5). Using qRT-PCR approach to study the expression patterns of these 5 miRNAs, in agreement with previous studies [5], we found that stress-associated miRNAs can display significantly lower expression under heat and drought.
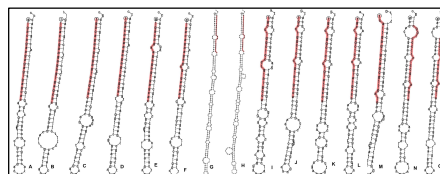

Figure 5. Secondary structure of the precursors of validated miRNAs: miR156a (A, B, C, D, E, F), miR159a (G, H), miR160a (I, J, K, L), miR397a (M) and miR408 (N, O). The miRNA sequences are highlighted in red. The minimal free energy for predicted secondary structures ranged from -37.60 to -92.8 kcal/mol, which falls within the range found for other plants (-8.5 to -180.8 kcal/mol; average -65.05 kcal/mol).

Conclusions

- Our analysis pipeline allowed to identify cassava miRNAs associated with abiotic stress and their targets genes.
- We identified over 800 potential microRNAs that represent a significant contribution to the improvement of the genomic resources for cassava

Acknowledgements

References

1. Meyers BC, Souret FF, Lu C, Green PJ: **Sweating the small stuff: microRNA discovery in plants**. *Curr Opin Biotechnol* 2006, **17**(2):139-146.
2. Ding Y, Tao Y, Zhu C: **Emerging roles of microRNAs in the mediation of drought stress response in plants**. *J Exp Bot* 2013.
3. Tuteja N: **Abscisic Acid and abiotic stress signaling**. *Plant Signal Behav* 2007, **2**(3):135-138.
4. Dai X, Cheng X, Li Y, Tang W, Han L: **Differential expression of gibberellin 20 oxidase gene induced by abiotic stresses in Zoysiagrass (Zoysia japonica)**. *Biologia* 2012, **67**(4):681-688
5. Zeng C, Wang W, Zheng Y, Chen X, Bo W, Song S, Zhang W, Peng M: **Conservation and divergence of microRNAs and their functions in Euphorbiaceous plants**. *Nucleic Acids Res* 2010, **38**(3):981-995.
6. Amiteye S, Corral JM, Sharbel TF: **Overview of the potential of microRNAs and their target gene detection for cassava(Manihot esculenta) improvement**. *Afr J Biotechnol* 2011, **10**(14):2562-2573